

---

# BNL Oracle database services status and future plans

---

Carlos Fernando Gamboa,  
John DeStefano,  
Dantong Yu

Grid Group, RACF Facility  
**Brookhaven National Lab, US**

Distributed Database Operations Workshop  
PIC Barcelona, Spain April 2008

# ATLAS Oracle DB services hosted at BNL

The US ATLAS production Database services at BNL are distributed among 2 independent oracle clusters

## RAC 1

### Dedicated to serve US ATLAS Conditions Database

-Besides serving BNL worker nodes, this database is also accessed by different Tier 2 and some Tier 3 site worker nodes distributed across the US. The US Muon calibration center also uses this cluster database via direct Oracle access or via Frontier.

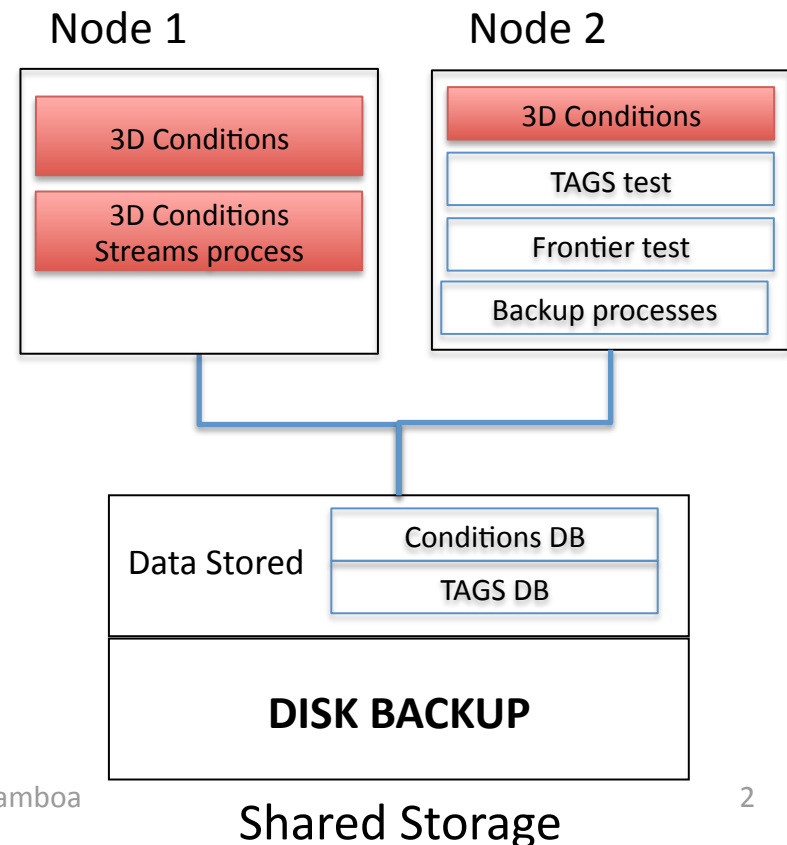
(References for different tests activities are included at the end of this presentation)

-Temporally hosts TAGS test data

-Database connection service most recently created and used by different Frontier tests activities.

-Streams replication process runs on a different node than the RMAN backup jobs.

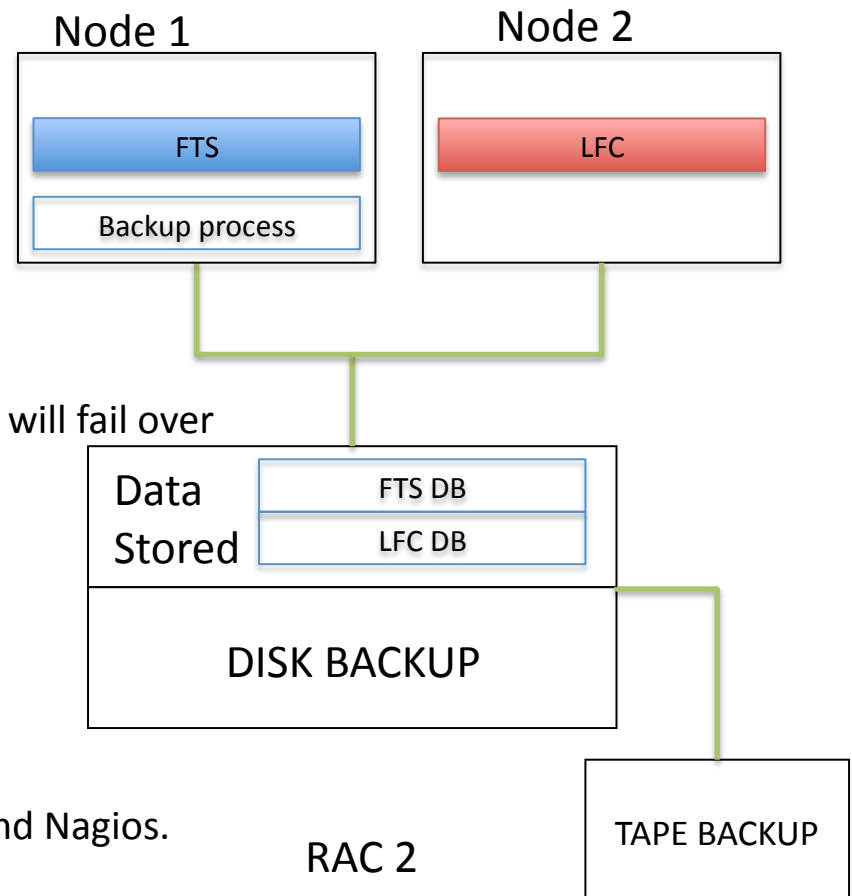
### Database service distribution RAC 1



# ATLAS Oracle DB services hosted at BNL

## RAC 2:

- Dedicated to host LFC and FTS data
- LFC recently integrated in US ATLAS production system (previously LRC used MySQL as backend)
- Each database service is distributed on only one node. In the case of failure, database services will fail over to the surviving node.
- Cluster inside BNL firewall
- TSM is enabled for tape backups besides the disks backups



BNL database services are monitor by OEM, Ganglia and Nagios.

# Outstanding BUGs found

---

## Bugs found

**Bug 7331323** - ORA-27302: FAILURE OCCURRED AT: SKGXPVFYMMTU

### Problem:

Caused an ASM instance crash one instance of the RAC 1 (Condition DB).

### Solution:

Special patch prepared by Oracle support. Not deployed in production since the bug has not occurred again.

**Bug. 6763804**, root.sh replaced by 2 lines that run rootadd.sh.

### Problem:

Not able to run root.sh script as required by CPU OCT 2008 patch if molecule 7155248 is present.

### Causing:

ORA-27369: job of type EXECUTABLE failed with exit code: ...

### Solution

Critical Patch Update October 2008 Database Known Issues Doc ID 735219.1

# Outstanding operational issues follow up by SR

---

## Conditions Database

### **SEPT 08. Service Request created SR(7203138.992).**

Apply process got stuck could not write data into the database. At this moment, backup process was running this could cause some resource contention at the OS level. Database user service was not affected.

### **Workaround:**

Isolate stream process on one node and run RMAN backups on different node. This issue has not happened since then.

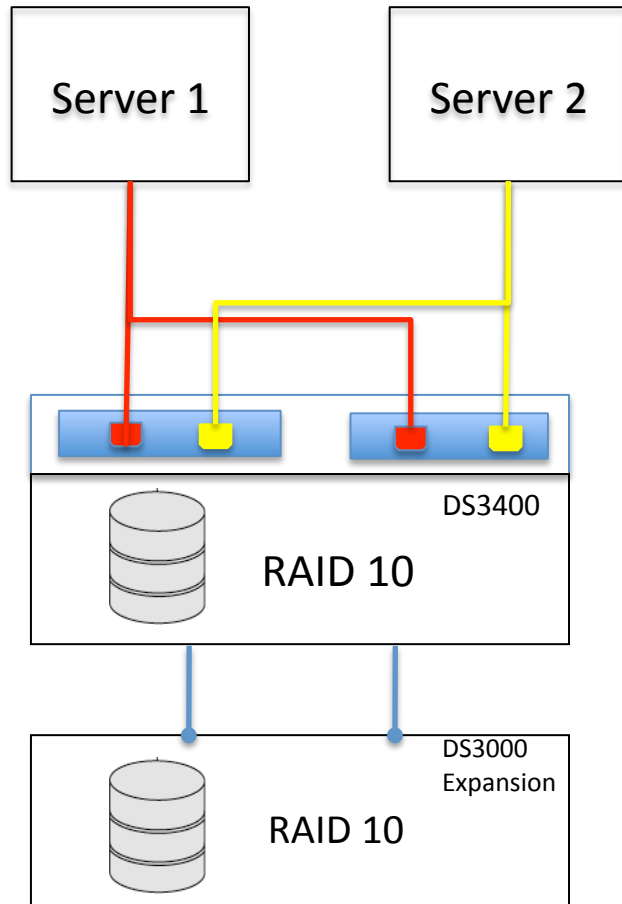
### **SR Status: Soft close**

Recently updated (03/18/09) by Oracle Engineer (previously on holidays), It is planned to move the backups to the same node to test whether the problem persists. In addition, trace files will be collected since the previous ones provided to support did not contain enough information to diagnose and understand the problem. This will be done as soon as current reprocessing activities ends.

Thanks to Eva for help following up on this issue.

# BNL ORACLE Cluster 2 nodes, DAS configuration

## general cluster hardware topology



### IBM 3550 Server description:

- 2 dual core 3GHz, 64 bits Architecture
- RAM 16GB

### Interconnectivity

Server to clients

- NIC 1000Gb/s.

Server to storage

- HBA QLogic 4Gb FC Dual-Port PCI-X
- 1M LC-LC Fibre Channel Cable

### Storage

IBM DS3400 FC dual controller

- 2 Hot Swap disk per enclosure
- 4 Gbps SW SFP Transceiver
- 12 SAS disks 15krpm and 10krpm, size 300 GB/disk

IBM DS3000 storage expansion

- 12 SAS disks 15krpm, size 300 GB/disk

## BNL Oracle services status for ATLAS

---

ATLAS production Oracle services hosted at BNL are distributed among 2 RAC clusters as:

RAC #	Oracle service	Nodes	Manufacture Model	Processor	Memory	HBA	NIC
1	Conditions DB	2	IBM 3550		16GB		1000Gb/s
2	FTS	1	IBM 3650	2 dual Core Intel Xenon Processor 5160 3GHz	8GB	QLogic 4Gb FC Dual-Port PCI-X	1000Gb/s
	LFC	1					

Production head nodes summary

# BNL Oracle services status for ATLAS

RAC #	Oracle service	Total RAW space	Total SPACE after RAID 10	Manufacture Model	Disk type, Speed, Size	Storage Controllers	Redundancy	IOPS per disk measured (ORION VERSION 11.1.0.7.0)
1	Conditions DB	6TB	2.8 TB	IBM DS3400, DS3000	SAS, 12 Disks 15K rpm 12 Disks 10K rpm 300GB	Dual FC controller 4 Gbps SW SFP Transceiver	Hot Swappable SAS disks Dual power supply	~200 IOPS / disk Measured with 5 LUNS RAID 1, 10 disks.
2	FTS LFC	6TB	2.8 TB	IBM DS3400, DS3000	SAS, 24 Disks 15K rpm, 300GB	Dual FC controller 4 Gbps SW SFP Transceiver	Hot Swappable SAS disks Dual power supply	

## Production storage summary



## BNL Oracle services status for ATLAS -General database distribution-

---

OS level 64 Bits	Database	Oracle Database Release	Data ASM disk Group	Backup ASM disk Group	SGA	Oracle ASMLibs
RHEL ES 2.6.9-78 release 4	Conditions DB				8 GB	
RHEL WS 2.6.9-78 release 4	FTS and LFC	10.2.0.4	1.4TB	1.4TB	4GB	2.0.2.1

# Future Plans

---

## Atlas Conditions Database

### Hardware

- Replace the head nodes with Quad core Nodes of 3GHz.
- To acquire 21.6 TB RAW storage, SAS disks of 450GB and 15krpm distributed among 1 DS3400 and 3 DS3000. Will use the same DAS technology.
- This modular technology allows to be integrated in an SAN configuration when required.

### Database Services

- TAGS database tests service will be allocated on a different test hardware
- Frontier test service

As of now, there are no procedures running on the database side for these tests. It is planned to include it to the FrONTier test infrastructure in future development efforts in coordination with CERN IT DB and Atlas DB group and frONTier developers folks.

The goal here is to minimize possible impact on the BNL production service (regular Cond.DB user service and T0->T1 database replication) while testing frONTier technology.

# Future Plans

---

## **Database configuration**

- Enable Jumbo Frames on internal network
- Decommission or modify existing password ATLAS\_COOL\_READER account
- Migration DB services to new hardware: Data guard or Transportable Table spaces.
- To test stand by databases

## **LFC and FTS**

- Migration from OS head nodes from WS to ES

# Round table topic

---

1. Is there any possibility to have an standard announcement format for interventions done at Tier 0 affecting TIER1 database service (replication), similar to the procedure used by Tier 1?

-This will facilitate the distribution of the intervention message to the affected user community by Tier 1 DBAs.

2. OS and database release, new hardware acquisition, new possibilities:

OS	DB
RHEL 4	11G
RHEL 5	11G
RHEL 5	10.2G

Are any of the above configurations running in production or being considered for deployment in the near future?

# References to relevant test activities involving database services at BNL

---

-Using FroNtier technology and the database access.

<http://indico.cern.ch/getFile.py/access?contribId=0&resId=0&materialId=slides&confId=53646>

And

[https://www.racf.bnl.gov/Facility/Frontier/Frontier\\_at\\_BNL\\_20090416.ppt](https://www.racf.bnl.gov/Facility/Frontier/Frontier_at_BNL_20090416.ppt)

-Involving the BNL batch system and the database direct access

<http://indico.cern.ch/getFile.py/access?subContId=0&contribId=4&resId=1&materialId=slides&confId=38539>

-Understanding client network access to the BNL database services

<http://indico.cern.ch/getFile.py/access?contribId=10&sessionId=2&resId=1&materialId=slides&confId=43856>

-TAGS tests, information can be found at

<https://twiki.cern.ch/twiki/bin/view/Atlas/EventTagInfrastructureAndTesting>